EXHIBIT 3
(PART 1 OF 2)

(54) Title: BREAST CANCER PROFILES AND METHODS OF USE THEREOF

3

(57) Abstract: This invention relates to the identification and use of gene expression profiles, or patterns, suitable for the identification of breast cancer patient populations with an inherited predisposition to breast and ovarian cancer. The gene expression patterns may be embodied in nucleic acid expression, protein expression, or other expression formats and may be used in the study and/or determination of optimal treatment, cancer prevention, patient and family identification, and other uses. The invention also pertains to the identification of patients with sporadic breast cancer, where a similar biology to that of hereditary breast cancer is caused by alternative mechanisms such as epigenetic modification of BRCAl or somatic mutation of other genes.

BREAST CANCER PROFILES AND METHODS OF USE THEREOF

FIELD OF THE INVENTION

[0001] This invention relates to the identification and use of gene expression profiles or patterns with clinical relevance to breast cancer.

GOVERNMENT LICENSE

[0002] The U.S. Government has a paid-up license in this invention and the right in limited circumstances to require the patent owner to license others on reasonable terms as provided for by the terms of National Institutes of Health (NIH) grant number P50 CA089018 awarded by the National Cancer Institute.

INTRODUCTION AND BACKGROUND OF THE INVENTION

Breast cancer and genetic risk

[0003] Approximately 212,920 new cases of invasive breast cancer, 61,980 in situ cases, and 40,970 deaths are expected to occur among US women in 2006 (Smigal C. et al. Trends in breast cancer by race and ethnicity: update 2006. CA: a Cancer Journal for Clinicians. 56(3):168-83, 2006.). Breast cancer is the leading cause of new cancers in women and comprises a third of all new cases. Breast cancer is the second leading cause of cancer mortality, accounting for 15% of the total deaths from cancer in women.

[0004] Breast cancer is a complex disease, resulting from an incompletely characterized interplay of genetic and environmental factors. About 5-10% of breast cancer is hereditary, i.e. due to the transmission of highly penetrant mutations in breast cancer predisposing genes. Within hereditary breast cancer families, mutation status is the overriding risk factor and genetic analysis can be used to clarify risk and guide medical management in a highly effective way. Genetic risk assessment consists of evaluating the pattern of cancers in the family, judging which of the known hereditary breast cancer syndromes fits the pattern, and pursuing genetic analysis.

[0005]   A specific genetic syndrome can be elucidated in about half of hereditary breast cancer families. Additional genes remain to be described (de Jong MM, Nolte IM, te Meerman GJ et al. Genes other than BRCA1 and BRCA2 involved in breast cancer susceptibility. J Med Genet 2002; 39(4):225-242). Risk-conferring alleles are conceptualized as high-penetrance genes with low prevalence [e.g. *BRCA1* and *BRCA2* [hereditary breast-ovarian cancer (HBOC) syndrome], *TP53* (Li-Fraumeni syndrome), *PTEN* (Cowden syndrome), *LKB1* (Peutz-Jeghers syndrome)] or low penetrance genes with high prevalence (possibly *CHEK2* (Meijers-Heijboer H, van den OA, Klijn J et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet 2002; 31(1):55-59), *ATM* (Thorstenson YR, Roxas A, Kroiss R et al. Contributions of ATM mutations to familial breast and ovarian cancer. Cancer Res 2003; 63(12):3325-3333) and the *TGFBR1*6A* allele (Kaklamani VG, Hou N, Bian Y et al. TGFBR1*6A and cancer risk: a meta-analysis of seven case-control studies. J Clin Oncol 2003; 21(17):3236-3243)

BRCA1 and BRCA2 gene function

[0006]   BRCA1 and BRCA2 encode very large proteins with 1,863 and 3,418 amino acids, respectively; each bears little homology to other known proteins or to each other. BRCA1 appears to play a role in numerous cellular functions including transcriptional regulation and influence of estrogen receptor activity, chromatin remodeling, DNA damage repair (homologous recombination and repair of transcription-coupled oxidation-induced DNA damage), centrosome duplication, cell growth, apoptosis, and cell cycle checkpoint control (Deng CX, Brodie SG. Roles of BRCA1 and its interacting proteins. Bioessays 2000; 22(8):728-737). BRCA1 contains an N-terminal RING domain that interacts with BARD1. Two BRCA1 C-terminal (BRCT) domains are present, which are found in proteins involved in DNA repair and control of the cell cycle. BRCA2 contains eight highly conserved BRC repeats of 30 to 40 residues in exon 11 which bind to RAD51, a key recombinational repair protein. After exposure of cells to DNA damage, BRCA1 relocalizes from nuclear foci to sites of DNA synthesis and becomes hyper-phosphorylated. BARD1, BRCA2, and RAD51 all relocalize with BRCA1 (Scully R, Livingston DM. In search of the tumour-suppressor functions of BRCA1 and BRCA2. *Nature* 2000; 408(6811):429-432). Germline mutations in BRCA1 are widely distributed throughout the gene (**FIG. 1**).

Clinical Significance of BRCA1 and BRCA2 mutations

[0007]   *BRCA1* and *BRCA2* mutations predispose female carriers to a high lifetime risk of breast cancer (>80%) and ovarian cancer (40-65% for *BRCA1* carriers and 20% for *BRCA2* carriers). The clinical features and management of HBOC syndrome have been reviewed (Lynch HT, Snyder CL, Lynch JF, Riley BD, Rubinstein WS. Hereditary breast-ovarian cancer at the bedside: role of the medical oncologist. *J Clin Oncol* 2003; 21(4):740-753). Average ages of breast and ovarian cancer onset are generally younger for *BRCA1* carriers than *BRCA2* carriers, but each can manifest as breast cancer in the twenties.

[0008]   Male breast cancer is seen in excess in *BRCA1* and *BRCA2* families, with about two thirds of positive cases involving *BRCA2* and one third involving *BRCA1* (Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol 2002*; 20(6):1480-1490). Lifetime risk of breast cancer is about 5-6% for male *BRCA1* and *BRCA2* carriers.

[0009]   Many effective cancer risk management strategies are available for *BRCA1* and *BRCA2* carriers, as well as for families with a high clinical suspicion of genetic predisposition (Scheuer L, Kauff N, Robson M et al. Outcome of preventive surgery and screening for breast and ovarian cancer in BRCA mutation carriers. *J Clin Oncol* 2002; 20(5):1260-1268). The chief value of genetic testing is to confirm the need for medical interventions, particularly those that are irreversible such as prophylactic mastectomy and prophylactic oophorectomy. As well, a true negative result (i.e. in the setting of a known familial mutation) obviates the need for aggressive surveillance and prevention measures, and provides reassurance to the person tested as well as to their offspring. Surveillance and management for HBOC syndrome includes consideration of chemoprevention of breast (e.g. tamoxifen) and ovarian (e.g. oral contraceptives) cancers, MRI surveillance for breast cancer, and early breast cancer surveillance (age 25 years) in at-risk female relatives(Scheuer L, Kauff N, Robson M, Kelly B, Barakat R,

Satagopan J et al. Outcome of preventive surgery and screening for breast and ovarian cancer in BRCA mutation carriers. *J Clin Oncol* 2002; 20(5):1260-1268).

Distinctive pathobiological features

[0010] *BRCA1* and *BRCA2* breast cancers have distinct biological features which differentiate them from sporadic or familial (non-*BRCA1/2*) breast cancers. At present, these distinguishing features are better recognized for *BRCA1* than *BRCA2* tumors.

[0011] A series of histopathologic and immunohistochemical (IHC) studies conducted by the Breast Cancer Linkage Consortium (BCLC) and other groups have revealed that *BRCA1* breast tumors, as compared with vs. age-matched sporadic breast cancers unselected for family history, are characterized by higher grade, higher mitotic counts, a greater degree of nuclear pleomorphism, less tubule formation, steroid receptor (ER/PR) negativity, HER-2 receptor negativity, lower p27(Kip1) protein levels, and cyclin E expression (Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. *Lancet* 1997; 349(9064):1505-1510; Lakhani SR, Jacquemier J, Sloane JP et al. Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J Natl Cancer Inst* 1998; 90(15):1138-1145; Chappuis PO, Kapusta L, Begin LR et al. Germline BRCA1/2 mutations and p27(Kip1) protein levels independently predict outcome after breast cancer. *J Clin Oncol* 2000; 18(24):4045-52; Lakhani SR, Van D, V, Jacquemier J et al. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol* 2002; 20(9):2310-2318).

[0012] *BRCA1* tumors share features of basal epithelial breast tumors such as cytokeratin (CK)5/6 expression and may largely overlap with this tumor subclass, based on IHC and gene expression profiling data. The basal/myoepithelial phenotype is seen in 2-18% of breast tumors, which are notable for IHC positivity for intermediate filaments e.g. CK5, CK14, usually high grade, with large central acellular zones comprising necrosis, tissue infarction, collagen, and

hyaline material, and ER, PR, HER-2 negative receptor status (Lakhani SR, Reis-Filho JS, Fulford L et al. Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. *Clin Cancer Res* 2005; 11(14):5175-5180). A model was developed for incorporating ER, CK14 and CK5/6 markers to select cases for *BRCA1* genetic testing. Marker status of ER negative and CK5/6 positive resulted in sensitivity = 56%, specificity = 97%, positive predictive value = 28% and negative predictive value = 99% with an area under the ROC curve = 0.77. The use of ER negative, CK14 and CK5/6 positive markers resulted in an area under the ROC curve = 0.87.

[0013]   *BRCA2* tumors are less distinctive, showing a higher overall grade as a result of exhibiting less tubule formation, and a higher proportion of continuous pushing margins, but are not significantly different with respect to mitoses, pleomorphism, and steroid receptor expression (Lakhani SR, Jacquemier J, Sloane JP et al. Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J Natl Cancer Inst* 1998; 90(15):1138-1145; Lakhani SR, Van D, V, Jacquemier J et al. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol* 2002; 20(9):2310-2318; Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. *Lancet* 1997; 349(9064):1505-1510). A comparison of *BRCA2* germline-mutated breast cancer vs. familial breast cancer using IHC of DNA repair proteins RAD51, RAD50, XRCC3, ATM, PCNA and CHEK2 showed that these tumors could be differentiated (Honrado E, Osorio A, Palacios J et al. Immunohistochemical expression of DNA repair proteins in familial breast cancer differentiate BRCA2-associated tumors. *J Clin Oncol* 2005; 23(30):7503-7511). CHEK2 expression was increased in *BRCA1* and *BRCA2* tumors vs. non-*BRCA1/2* and sporadic tumors. *BRCA2* breast tumors showed absent RAD51 nuclear expression and had cytoplasmic RAD51 expression. The results were validated with a new series of patient cases and a multivariate model was developed with RAD51 and CHEK2 that distinguishes *BRCA2* from non-*BRCA1/2* tumors with an estimated probability of ≥ 76%.

[0014]   *BRCA1* and *BRCA2* breast tumors are more likely to overexpress p53 and more commonly harbor somatic mutations in the *TP53* gene with an altered mutational spectrum,

suggesting that that impaired DNA repair function may play a central role in molecular pathogenesis (Greenblatt MS, Chappuis PO, Bond JP, Hamel N, Foulkes WD. TP53 mutations in breast cancer associated with BRCA1 or BRCA2 germ-line mutations: distinctive spectrum and structural distribution. *Cancer Res* 2001; 61(10):4092-4097).

Gene expression profiling of BRCA breast tumors

[0015]   Gene expression profiling (GEP) has become an important tool for the comprehensive analysis of gene expression in diverse biological samples and has emerged as a means for refining the taxonomy of cancers. This method may help to clarify prognosis, optimize treatment, elucidate molecular progression pathways, and lead to the development of new cancer therapeutics tailored to the underlying etiology. The independent prognostic value of gene expression signatures in early stage breast cancer has already led to the development of clinical tests and engendered clinical trials ('t Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002; 415(6871):530-536; Paik S, Shak S, Tang G et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004; 351(27):2817-2826: Esteva FJ, Sahin AA, Cristofanilli M et al. Prognostic role of a multigene reverse transcriptase-PCR assay in patients with node-negative breast cancer not receiving adjuvant systemic therapy. *Clin Cancer Res* 2005; 11(9):3315-3319; Tuma RS. A big trial for a new technology: TransBIG Project takes microarrays into clinical trials. *J Natl Cancer Inst* 2004; 96(9):648-649).

[0016]   Gene expression patterns have been used to discern "molecular portraits" of breast tumors (Perou CM, Sorlie T, Eisen MB et al. Molecular portraits of human breast tumours. *Nature* 2000; 406(6797):747-752). Breast tumor subtypes distinguished by DNA microarrays appear to represent distinct biological entities: luminal subtypes A and B, ERBB2+ subtype, basal subtype, and normal breast-like subtype (Perou CM, Sorlie T, Eisen MB et al. Molecular portraits of human breast tumours. *Nature* 2000; 406(6797):747-752).

[0017]   Genomic methods have been employed to "bin" hereditary tumors. For example, comparative genomic hybridization of non-*BRCA1/2* hereditary breast tumors was used to guide the mapping of additional susceptibility genes (Kainu T, Juo SH, Desper R et al. Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer

susceptibility locus. *Proc Natl Acad Sci* U S A 2000; 97(17):9603-9608), and distinguished *BRCA1*-mutated from sporadic breast tumors with an accuracy of 84% (Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. *Cancer Res* 2002; 62(23):7110-7117).

[0018]   Notably, GEP can accurately distinguish *BRCA1*, *BRCA2*, and sporadic breast tumors (Hedenfalk I, Duggan D, Chen Y et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001; 344(8):539-548; 't Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415(6871):530-536). Review of 176 differentially expressed genes revealed a common theme in *BRCA1* mutated samples, involving the coordinated transcriptional activation of two major cellular processes, DNA repair and apoptosis (Hedenfalk I, Duggan D, Chen Y et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001; 344(8):539-548). A *BRCA1* signature was also discerned in a study which used GEP to identify "poor prognosis" signatures in breast tumors from young women with node-negative disease ('t Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415(6871):530-536). Using an optimal set of 100 *BRCA1* reporter genes, the investigators were able to distinguish *BRCA1* from sporadic ER negative breast cancers with an accuracy of 95%. "Misclassified" sporadic tumors had decreased *BRCA1* expression and promoter hypermethylation, reflecting a common biology between germline- and somatically inactivated tumors and showing the centrality of *BRCA1* in determining the molecular phenotype (Hedenfalk I, Duggan D, Chen Y et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001; 344(8):539-548; 't Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415(6871):530-536). All of the *BRCA1* tumors fell within the basal subgroup, indicative of a distinctive biology associated with a poor prognosis (Sorlie T, Tibshirani R, Parker J et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci* U S A 2003; 100(14):8418-8423). *BRCA2* tumors fell within the luminal A subtype.

[0019]   Gene expression profiling studies demonstrate that a highly penetrant susceptibility gene can markedly influence the molecular phenotype, histology, and prognosis of the resulting

breast tumor. Moreover, the molecular phenotype can be examined to gain insight into the specific cellular pathways that have been disrupted (Hedenfalk I, Duggan D, Chen Y et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001; 344(8):539-548).

Prognosis

[0020]   *BRCA1* breast tumors show a poorer survival rate as compared with matched sporadic and *BRCA2* controls in some, but not all studies(Robson ME, Boyd J, Borgen PI, Cody HS3. Hereditary breast cancer. *Curr Probl Surg* 2001; 38(6):387-480; Robson ME, Chappuis PO, Satagopan J et al. A combined analysis of outcome following breast cancer: differences in survival based on BRCA1/BRCA2 mutation status and administration of adjuvant treatment. Breast Cancer Res 2004; 6(1):R8-R17). The survival disadvantage in *BRCA1* carriers may disappear if patients with small, node-negative grade 3 tumors are treated with chemotherapy (Robson ME, Boyd J, Borgen PI, Cody HS3. Hereditary breast cancer. *Curr Probl Surg* 2001; 38(6):387-480; Robson ME, Chappuis PO, Satagopan J et al. A combined analysis of outcome following breast cancer: differences in survival based on BRCA1/BRCA2 mutation status and administration of adjuvant treatment. *Breast Cancer Res* 2004; 6(1):R8-R17; Evans DG, Howell A. Are B. *Breast Cancer Res* 2004; 6(1):E7).

[0021]   Interestingly, the possibility of worse prognosis with node-negative tumors is paralleled by a large study showing disruption of the expected positive correlation between breast tumor size and lymph node status in BRCA1 breast cancers (Foulkes WD, Metcalfe K, Hanna W et al. Disruption of the expected positive correlation between breast tumor size and lymph node status in BRCA1-related breast carcinoma. *Cancer* 2003; 98(8):1569-1577). Among 1555 women with invasive breast cancers diagnosed between 1975-1997 comprised of 276 *BRCA1* mutation carriers, 136 BRCA2 carriers, and 1143 women without a known mutation (208 *BRCA1/BRCA2* noncarriers and 935 untested women), a highly significant positive correlation was found, as expected, between tumor size and the frequency of positive axillary lymph nodes among BRCA1/BRCA2 noncarriers, BRCA2 carriers, untested women (overall $P < 0.0001$ for each). Notably however, no clear correlation was found between tumor size and positive lymph node status in BRCA1 carriers (overall $P = 0.20$). If this relationship seen in most tumors is lost, then it stands to reason that small, lymph-node negative BRCA1 tumors may

nonetheless carry an adverse prognosis. If such tumors are amenable to treatment, then a more aggressive approach (i.e. chemotherapy for small, node-negative tumors) might be warranted.

[0022]   Clinical survival studies are intriguing, particularly when placed into context with characteristics of basal tumors. Basal tumors are characteristically large and express low levels of ER, HER2, and p27Kip1, high levels of cyclin E, with nuclear p53 and intratumoral vascular nests (also referred to as glomeruloid-microvascular-proliferation or GMP) [Foulkes WD, Brunet JS, Stefansson IM et al. The prognostic implication of the basal-like (cyclin E high/p27 low/p53+/glomeruloid-microvascular-proliferation+) phenotype of BRCA1-related breast cancer. *Cancer Res* 2004; 64(3):830-835]. All of these factors are associated with a poor outcome in univariate analyses, and tumor markers most closely linked to the basal phenotype (p53, p27Kip1, cyclin E, and GMP) are independent predictors of poor outcome. Taken together, these data suggest that much of the inferior survival experienced by *BRCA1* carriers with breast cancer—particularly those with lymph node-negative disease—may be attributable to the basal epithelial phenotype of these cancers (Foulkes WD, Brunet JS, Stefansson IM et al. The prognostic implication of the basal-like (cyclin E high/p27 low/p53+/glomeruloid-microvascular-proliferation+) phenotype of BRCA1-related breast cancer. *Cancer Res* 2004; 64(3):830-835).

[0023]   A study examining the efficacy of neo-adjuvant chemotherapy found a better clinical response rate in *BRCA1/2* carriers than in non-carriers. The probability of achieving a complete response in *BRCA1/2* carriers seems to be independent of stage, suggesting that if inferior survival is a characteristic of *BRCA* tumors, it may be amenable to treatment using chemotherapy (Chappuis PO et al. A significant response to neoadjuvant chemotherapy in BRCA1/2 related breast cancer. *J Med Genet* 2002; 39(8):608-610).

Treatment tailored to genotype- chemotherapy

[0024]   Foulkes has recently reviewed the *in vitro* and *in vivo* data on chemosensitivity of *BRCA1/2* breast tumors (Foulkes WD. BRCA1 and BRCA2: chemosensitivity, treatment outcomes and prognosis. *Fam Cancer* 2006; 5(2):135-142). Questions that are ripe for inquiry include whether platinum-based therapies are more effective than taxanes for *BRCA1/2* carriers. While anthracycline treatment has shown good results in the clinical setting, the data are not definitive and the in vitro data are less encouraging. Randomized, controlled clinical trials will

be required to answer these questions. This raises the logistical issues involved in elucidating the mutation status of *BRCA1/2* carriers at the time of breast cancer diagnosis, so as to enable treatment studies.

Targeting of therapy to underlying biology

[0025]   BRCA1 and BRCA2 are important for DNA double strand (DS) break repair by homologous recombination. Poly(ADP-ribose) polymerase (PARP) is an enzyme involved in base excision repair, a key pathway in the repair of DNA single strand (SS) breaks. BRCA1 or BRCA2 dysfunction profoundly sensitizes cells to the inhibition of PARP enzymatic activity, resulting in chromosomal instability, cell cycle arrest and subsequent apoptosis. This seems to be because the inhibition of PARP leads to the persistence of DNA lesions normally repaired by homologous recombination. These results illustrate how different pathways cooperate to repair damage, and suggest that the targeted inhibition of particular DNA repair pathways may allow the design of specific and less toxic therapies for cancer (Farmer H, McCabe N, Lord CJ et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* 2005; 434(7035):917-921).

[0026]   PARP1 facilitates DNA repair by binding to DNA breaks and attracting DNA repair proteins to the site of damage. Nevertheless, PARP-/- mice are viable, fertile and do not develop early onset tumours. PARP inhibitors trigger g-H2AX and RAD51 foci formation. Bryant et al. (Bryant HE, Schultz N, Thomas HD et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature 2005; 434(7035):913-917) propose that, in the absence of PARP1, spontaneous SS breaks collapse replication forks and trigger homologous recombination for repair. Furthermore, they show that BRCA2-deficient cells, as a result of their deficiency in homologous recombination, are acutely sensitive to PARP inhibitors, presumably because resultant collapsed replication forks are no longer repaired. Thus, PARP1 activity is essential in homologous recombination-deficient BRCA2 mutant cells. They exploit this requirement in order to kill BRCA2-deficient tumours by PARP inhibition alone. Treatment with PARP inhibitors is likely to be highly tumour specific, because only the tumours (which are BRCA2-/-) in BRCA2+/- patients are defective in homologous recombination. The use of an

inhibitor of a DNA repair enzyme alone to selectively kill a tumour, in the absence of an exogenous DNA-damaging agent, represents a new concept in cancer treatment.

SUMMARY OF THE INVENTION

[0027]   The method described herein relates to the identification of gene expression profiles or patterns of certain genes linked to the function of a gene known as BRCA1 in breast or ovarian cancer. The method is useful for the identification of individuals with hereditary predisposition to breast and ovarian cancer, such that appropriate cancer prevention or treatment options may be implemented.

[0028]   In particular, the method described here relates to detecting the presence of hereditary mutations in BRCA1 or the BRCA1 pathway which disrupt downstream gene expression. Preferably, the method is applied to archival breast or ovarian tissue samples which have been formalin-fixed and embedded in paraffin (FFPE). The mRNA samples in such FFPE tissues are degraded and may not be useful for conventional DNA arrays. Thus, in one embodiment of the method described here, the gene profiles are established using a DNA array designed to amplify mRNA signal from degraded samples embedded in paraffin following formalin fixation. Most preferably, the DNA array is an Illumina DASL array (a cDNA-mediated annealing, selection, extension, and ligation assay) or other array specifically designed for degraded mRNA samples.

[0029]   The method of using gene expression profiling to detect the presence of functional BRCA1 mutations is independent of the estrogen-receptor (ER) status of the tissue sample being analyzed. Thus, ER-positive tissue samples from breast tissue specimens will generate data that are similar to ER-negative samples for purposes of BRCA1 analysis using the method described herein.

[0030]   In one embodiment of the method described here, the gene expression profile is established by selecting at least 10 genes from a group of 128 candidates and analyzing the mRNA expression using a DNA array. In one especially preferred embodiment, 13 genes from the larger group of 128 genes are profiled to distinguish sporadic BRCA1 mutations from

hereditary mutations. In another embodiment of this method, at least 2 genes from the subset of 13 genes are selected for analysis of mRNA expression in FFPE breast or ovarian tissue.

[0031]   In a further embodiment of this method, the sensitivity of the method in detecting hereditary BRCA1 mutations in FFPE tissues is greater than or equal to 70%. In a still further embodiment of this method, the sensitivity of this method in detecting hereditary BRCA1 mutations in FFPE tissue is greater than or equal to 80%.

[0032]   In another aspect of the described method, the specificity of the method in distinguishing between sporadic and hereditary BRCA1 mutations is greater than or equal to 70%. A still further aspect of this method provides for distinguishing between sporadic and hereditary BRCA1 mutations with a specificity greater than or equal to 80%.

[0033]   In an alternative embodiment, the method can be used to detect loss of *BRCA1* function in cancers that are the result of somatic pathways including genes that are upstream or downstream of BRCA1 in a biological pathway. These upstream or downstream genes regulate BRCA1 function and may decrease the activity of BRCA1, resulting in a gene profile or pattern that is similar to when the mutations occur in BRCA1 itself.

BRIEF DESCRIPTION OF THE FIGURES

[0034]   **FIG.1** is a graph indicating that germline BRCA1 mutations in breast cancer specimens that this invention is capable of identifying are widely distributed across the BRCA1 gene.

[0035]   **FIG. 2** illustrates gene expression profiles of 14 probes (for 13 genes) that are differentially expressed in *BRCA1*-mutated breast tumors in comparison with sporadic breast tumors.

[0036]   **FIG. 3** Plot of *BRCA1* expression levels vs. methylation status of *BRCA1* promoter. Among sporadic breast cancers, *BRCA1* expression levels were inversely correlated with methylation of the *BRCA1* promoter ($P<0.01$).

[0037]   **FIG. 4** is a graph showing the relationship of RNA quality vs. age of archival sample. High Ct value reflects poorer RNA quality. Age of archival material is not predictive of sample quality. The oldest sample (39 years) demonstrates one of the highest quality RNAs.

[0038]   **FIG. 5** is a chart showing the distribution of genes selected for the custom array into several functional categories, with particular weighting towards transcriptional regulation, cell cycle control, and DNA repair.

[0039]   **FIG. 6** is a graph showing qPCR data for MAGEA4 mRNA expression comparing BRCA1 mutated samples (ER positive and ER negative) versus sporadic (ER positive and ER negative) samples.

[0040]   **FIG. 7** is a graph showing qPCR data for SPIB mRNA expression comparing BRCA1 mutated samples (ER positive and ER negative) versus sporadic (ER positive and ER negative) samples.

[0041]   **FIG. 8** is a graph showing qPCR data for BRCA2 mRNA expression comparing BRCA1 mutated samples (ER positive and ER negative) versus sporadic (ER positive and ER negative) samples.

[0042]   **FIG. 9** is a side-by-side comparison of graphs generated using the MAGE A4 qPCR data in FIG. 6 compared to data generated using a DASL array.

[0043]   **FIG. 10** is a side-by-side comparison of graphs generated using the SPIB qPCR data in FIG. 7 compared to data generated using a DASL array.

[0044]   **FIG. 11** is a side-by-side comparison of graphs generated using the BRCA2 qPCR data in FIG. 8 compared to data generated using a DASL array.

[0045]   TABLE 1 is a chart depicting the gene ontology classification of 120 non-control genes selected for the custom array.

[0046]   TABLE 2 is a chart depicting genes in the *BRCA1* classifier which are implicated in stem cell biology.

[0047]   TABLE 3 is a chart listing the gene symbols and descriptions of the genes in the 128-gene array.

[0048]   TABLE 4 is a chart listing the various database identifiers for the genes in the 128-gene array.

[0049]   TABLE 5 is a chart listing the 13-gene *BRCA1* classifier selected from the broader 128-gene array.

<u>DETAILED DESCRIPTION OF VARIOUS EMBODIMENTS</u>

[0050]   The present invention relates to the use of gene expression profiles (alternatively described as "profiles" or "signatures") which are clinically relevant to breast cancer (for background purposes, please see Erlander et al., U.S. Patent Application Publication US 2005/0095607, hereby incorporated by reference in its entirety). In particular, the identities of genes which are correlated with hereditary breast cancer due to inherited mutations in the BRCA1 gene are provided. The gene expression profiles, whether embodied in nucleic acid expression, protein expression, or other expression formats, may be used to identify breast tumors with non-functional BRCA1 genes. Non-functioning of the BRCA1 gene may be due to germline mutations in the BRCA1 gene and/or acquired loss of function in the BRCA1 gene. Identification of breast tumors with BRCA1 loss of function is not solely dependent on analysis of the BRCA1 gene, but relies on analysis of additional genes and their pattern of expression. The methods described herein may be used to define the functional and clinical significance of variants in the BRCA1 gene, including missense mutations, thereby categorizing variants as disease-causing or clinically benign. The methods relate to the analysis of breast tumors

including but not limited to archival tumor materials that are formalin-fixed and paraffin-embedded.

[0051]   The identification of BRCA1 and BRCA2 germline mutation carriers is done primarily for the purpose of cancer risk management. A wide variety of clinically effective early detection and prevention strategies are available to female carriers. In order to take full advantage of these clinical treatments, the mutation status must be identified. This is currently done using direct analysis of the BRCA1 and BRCA2 genes mainly through DNA sequencing. DNA expression analysis may also be conducted in conjunction with protein expression analysis by various methods. For example, mRNA levels for genes relating to BRCA1 may be co-analyzed with protein expression levels by employing such methods as immunohistochemistry (IHC).

[0052]   While one gene may be accurate to discriminate BRCA1 loss of function, more genes will tend to provide more accuracy. It is contemplated that the method disclosed herein use multiple genes disclosed in **TABLE 3**.

[0053]   As used herein, these terms shall be defined as follows:

[0054]   "Gene expression profile or pattern" shall refer to the mRNA expression of certain genes that are either over-expressed or under-expressed when *BRCA1* is mutated in comparison to a normal or functional *BRCA1* gene. Combining the data of at least 2 individual genes constitutes a profile or pattern which can be used to assess the functional status of the *BRCA1* gene.

[0055]   "Array" or "microarray" refers to a substantially 2-dimensional arrangement of polynucleotides specifically placed on a solid support such as glass, plastic, beads, or other synthetic material in such a way that the location of the polynucleotide on the array is fixed in relation to other polynucleotides on the same array, thus allowing for the user to correlate data from an assay using the array with specific polynucleotides of known locations on the array. An array may allow for enzymatic reactions on the surface of the support such as annealing of primers or other exogenous polynucleotides, extension of said polynucleotides, or ligating of added nucleotides or polynucleotides.

[0056]   "Mutation" refers to a substitution, deletion, or addition of a nucleotide or nucleotides to the wildtype sequence of a gene as identified herein. Mutations may be both "functional" or "non-functional", i.e. a "silent" mutation may occur where a substitution mutation results in the same amino acid sequence for the translated gene product, or a mutation may result in an amino acid sequence change which renders the translated protein non-functional. A mutation in a gene of interest may result in genes further downstream in a biological cascade to change mRNA expression either positively or negatively.

[0057]   "Specificity" of the method described herein refers to the percent accuracy with which the method distinguishes gene profiles of sporadic tumors versus tumors arising from hereditary mutations.

[0058]   "Sensitivity" of the method described herein refers to the percent accuracy with which the method detects hereditary *BRCA1* mutations in tissue samples containing functional mutations.

[0059]   "Distinguishes" as used herein denotes the usefulness of an assay in categorizing certain tumors or mutations as either sporadic or hereditary. The gene profiling as described herein "distinguishes" between these alternatives when the statistical probability that the change in mRNA expression level for a particular gene above or below baseline levels is due to chance alone is less than 1 percent by Chi Square analysis or 5 percent by Student's t-test.

[0060]   "Estrogen Receptor status" refers to the presence or absence of estrogen receptor on the surface of cells in the tissue or tumor sample being analyzed.

[0061]   "Sporadic" refers to mutations or tumors arising in breast or ovarian tissues caused by environmental or other factors that does not include those highly penetrant mutations in breast cancer predisposing genes inherited from either or both parents of the individual. Sporadic tumors may include low-penetrance genes inherited from either or both parents of the individual. Sporadic tumors may also arise as a result of DNA methylation of the *BRCA1* gene or promoter or other epigenetic mechanisms.

[0062]   "Hereditary" refers to those mutations present from the earliest stages of development of the individual or organism which were inherited from either or both parents, or arose de novo in an individual and can be transmitted to offspring in subsequent generations.

[0063]   "BRCA1" refers to the DNA, mRNA, or translated protein of the gene identified herein as UG Rep Acc NM_007295, LLID 672, and physically chromosomally located at cytoband 17q21.

[0064]   "Formalin-fixed paraffin-embedded (FFPE)" refers to archival tissue samples which are initially fixed in formalin prior to being embedded in paraffin wax and allowed to cool into solids, whereby they can be maintained at room temperature for extended periods of time before being analyzed by the procedures of the method described herein including mRNA extraction and microarray analysis for the purposes of gene profiling.

Limitations of family history analysis to identify at-risk individuals

[0065]   Identification of patients for BRCA1 and BRCA2 gene sequencing relies heavily on a clinician taking a detailed family history of cancer, then acting on this information by referring the patient for genetic counseling and genetic testing. There are several shortcomings to this approach, which are unrelated to the methods for BRCA1 and BRCA2 gene mutations. While the sensitivity of mutation analysis is high, approximately 90%, the vast majority of BRCA1 and BRCA2 mutation carriers are clinically unrecognized. These are detailed as follows.

[0066]   Clinical recognition of BRCA1 and BRCA2 germline mutation status relies on family history but family history is indicative of an underlying mutation in 50% or fewer cases of female breast cancer. This is evidenced by population-based studies of women with incident breast cancer cases (Peto J et al. *J Natl Cancer Inst* 1999, 91:943–9; Hopper JL et al. *Cancer Epidemiol Biomarkers Prev* 1999, 8: 41–7; King MC et al., *Science* 2003, 302:643–6; de Sanjose S et al., *Int J Cancer* 2003, 106:588–93; Warlam-Rodenhuis CC et al. *Eur J Cancer* 2005; 41:1409-15, 2005). In all reported studies except that of King et al., women were stratified according to young ages at breast cancer diagnosis, further indicating the limitations of family history in identifying at-risk women.

[0067]   The rate of carrier identification in the United States is 10% or less, using the following parameters: 180,000 breast cancer cases diagnosed each year, of which 5% are due to BRCA1 or BRCA2 mutations = 9000 BRCA1 or BRCA2 related breast cancers per year; 10 years during which BRCA1 and BRCA2 DNA sequencing has been clinically available = 90,000 BRCA1 and BRCA2 related breast cancers over the past 10 years; 13055 carriers have been identified (Martin et al., Annual Meeting of the American Society for Human Genetics, Oct. 10, 2006, abstract #371), but this includes women with and without a diagnosis of breast cancer). According to report of the first 10,000 cases (Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol* 2002; 20(6):1480-1490), fewer than half (4843) of women who had gene testing had breast cancer. Thus approximately half of the 13055 carriers (2006) would be expected to have breast cancer meaning that about 6500 of a potential 90,000 BRCA1 and BRCA2 related breast cancers over the past 10 years have been identified, only 7%.

[0068]   Direct tumor analysis as described herein can be used as a method to identify cases which are otherwise indiscernible using family history. Identification of carriers would not be limited by clinical parameters such as young age at breast cancer diagnosis.

Limitation of sample availability for gene testing to living individuals

[0069]   Gene testing of BRCA1 and BRCA2 using comprehensive mutation analysis is currently limited to DNA samples derived from blood specimens. In the vast majority of cases this means that testing is restricted to living persons (exceptions include people whose DNA was banked prior to their death, or Ashkenazi Jewish individuals whose archival tumor specimens can be subjected to limited DNA analysis). This severely limits the clinical utility of gene testing. This is because comprehensive mutation analysis within families is ideally first performed on an individual who has had breast or ovarian cancer (or both). When testing is performed in this manner, a positive result is more likely to be obtained (given autosomal dominant inheritance, unaffected individuals are about half as likely to test positive). Positive results confirm the hereditary condition in the family and also provide for clear cut results in

relatives subsequently testing (true positive or true negative). Because the test sensitivity of BRCA1 and BRCA2 comprehensive mutation analysis is less than 100% and because genes other than BRCA1 and BRCA2 cause breast cancer, a negative result is uninformative until a positive result has been observed within a family. Thus, interpretable results often hinge on the availability of a blood specimen from an affected individual. However such individuals have often died due to the aggressive nature of breast and ovarian cancer. By extending the analysis to archival tumor specimens, the availability of genetic testing in families is enhanced.

Targeting gene testing to those likely to have mutations; reducing cost of gene testing

[0070]   The sensitivity of gene testing in women with breast cancer using traditional comprehensive mutation analysis methods can be enhanced by using gene expression profiling of breast tumors. According to Frank et al. (Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol* 2002; 20(6):1480-14902002) only 20% of women with a history of breast cancer who undergo comprehensive mutation analysis have identifiable BRCA1 or BRCA2 mutations. This is a costly approach since each comprehensive analysis costs about $3000, with a corresponding cost of $15,000 to identify a single mutation carrier. By using gene expression profiling as a pre-screen, comprehensive mutation analysis could then be targeted to women whose breast cancers have a high (80-90%) likelihood of being due to an underlying germline mutation. Furthermore, comprehensive mutation analysis could be restricted to the gene in question (e.g. BRCA1 only as opposed to both BRCA1 and BRCA2), further reducing costs. Once a mutation is detected, then genetic testing in blood relatives is enhanced in that the cost is less (~$400) because single-site analysis can be done, and the accuracy approaches 100% for both positive and negative results.

Utility of identifying patients whose breast tumors are not due to BRCA1 mutations

[0071]   Patients whose tumors have a BRCA1 gene expression pattern can be tested for BRCA1 only. If a mutation is not identified, these patients can become the subject of additional research studies. Underlying possibilities include a missed mutation; using specimens from these patients new methods for detecting underlying mutations can be developed. Another possibility is that other inherited genes are involved which are likely either upstream or

downstream of the BRCA1 gene; these patients and their families can become the subject of linkage analysis and other methods to identify novel cancer predisposing genes. Alternatively, these patients may have somatic (acquired) mutations, the biology of which can be further explored given implications for targeted treatment and possibly worse prognosis.

Functional assay; clarification of variants

[0072] Genetic testing of most hereditary cancer genes is hampered by the high prevalence of variants of uncertain significance. These are DNA sequence variants which may or may not compromise the function of the gene, but for which there is insufficient information to characterize their clinical function. In the case of BRCA1 and BRCA2, variants of uncertain clinical significance comprise about 7% of test results.

[0073] Epidemiological and biological criteria can be applied to distinguish functional from benign variants with some success (Deffenbaugh AM, Frank TS, Hoffman M, Cannon-Albright L, Neuhausen SL. Characterization of common BRCA1 and BRCA2 variants. *Genet Test* 2002; 6(2):119-121). For example, the prevalence of each variant in a control population, co-segregation of the variant with cancer within families, location of the variant within the gene, functional assays, demonstration of abnormal mRNA transcript processing, type of the amino acid substitution and degree of conservation among species (Fleming MA, Potter JD, Ramirez CJ, Ostrander GK, Ostrander EA. Understanding missense mutations in the BRCA1 gene: an evolutionary approach. *Proc Natl Acad Sci USA* 2003; 100(3):1151-1156) provide clues as to whether the mutation is deleterious (Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol* 2002; 20(6):1480-1490). Genetic counseling for variants of uncertain clinical significance is a commonly encountered and highly problematic issue (Petrucelli N, Lazebnik N, Huelsman KM, Lazebnik RS. Clinical interpretation and recommendations for patients with a variant of uncertain significance in BRCA1 or BRCA2: a survey of genetic counseling practice. *Genet Test* 2002; 6(2):107-113), which can possibly lead to the inappropriate use of medical interventions such as prophylactic surgery (Lynch HT, Snyder CL, Lynch JF, Riley BD, Rubinstein WS. Hereditary breast-ovarian cancer at the bedside: role of the medical oncologist. *J Clin Oncol* 2003; 21(4):740-753).